Cover Page

1) Title of the paper:

Quality metrics can help the expert during neurological clinical trials

2) authors' affiliation and address:

IRCCyN-IVC, (UMR CNRS 6597), Polytech' Nantes
Rue Christian Pauc, La Chantrerie, 44306 NANTES, France.
Tel : 02.40.68.30.52
Fax : 02.40.68.32.32

3) e_mail address:

Florent.Autrusseau@univ-nantes.fr

5) bibtex entry:

```
@inproceedings{MI2016,
  author = {L. Mahé, F. Autrusseau, H. Desal, J. Guedon, H. Der
Sarkissian, Y. Le Teurnier, S. Davila},
  title = {Quality metrics can help the expert during
neurological clinical trials},
  booktitle = {Medical Imaging},
  year = {2016}
}
```

Quality metrics can help the expert during neurological clinical trials

L. Mahé[a,b], F. Autrusseau[a], H. Desal[c], J. Guedon[a], H. Der Sarkissian[b], Y. Le Teurnier[c], S. Davila[b]
[a]Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, FRANCE; [b]KEOSYS, Saint-Herblain, FRANCE; [c]Hôpital Nord Laennec, service de neuroradiologie diagnostique et interventionnelle, Nantes, FRANCE

ABSTRACT

Carotid surgery is a frequent act corresponding to 15 to 20 thousands operations per year in France. Cerebral perfusion has to be tracked before and after carotid surgery. In this paper, a diagnosis support using quality metrics is proposed to detect vascular lesions on MR images. Our key stake is to provide a detection tool mimicking the human visual system behavior during the visual inspection. Relevant Human Visual System (HVS) properties should be integrated in our lesion detection method, which must be robust to common distortions in medical images. Our goal is twofold: to help the neuroradiologist to perform its task better and faster but also to provide a way to reduce the risk of bias in image analysis. Objective quality metrics (OQM) are methods whose goal is to predict the perceived quality. In this work, we use Objective Quality Metrics to detect perceivable differences between pairs of images.

Keywords: objectives quality metrics, human visual system, lesion detection, MRI, neurology.

1. INTRODUCTION

Carotid endarterectomy is a frequent act corresponding to 15 to 20 thousands operations per year in France. This surgery aims to prevent stroke in carotid artery disease patients. This condition is caused by an inadequate blood flow through the carotid arteries, related to a plaque buildup in these arteries. Thus, the supply of oxygen and nutrient to the brain tissue is reduced or cut off. Today, a few technical innovations offer a follow-up of the arterial oxygen flow during the carotid endarterectomy. The monitoring of cerebral hemodynamic during surgery may greatly reduce the brain-damaged occurring after carotid clamping stage. However, these kinds of medical device have to demonstrate their potential abilities. This is the core of the EMOCAR clinical trial.

EMOCAR is a multi-site clinical trial which has already enlisted more than 800 patients in different French hospitals. Each patient gets a first cerebral MRI before carotid surgery and a second cerebral MRI two or three days after surgery. The neuroradiologist has to perform a visual comparison between the two volumes looking for the presence of new lesions or shape distortion into the vascular tree on a regular working station. Therefore, the prime figure of merit (FOM) is the number of new cerebral ischemic lesions seen on post-surgical MRI, as seen in Figure 1. This perceptual lesion detection must be differentiated from the lesion detection performed by some other ways such as segmentation process. This task is directly related to the radiologist's expertise and can be jeopardized by various external factors such the image quality but, also, the image acquisition and analysis.
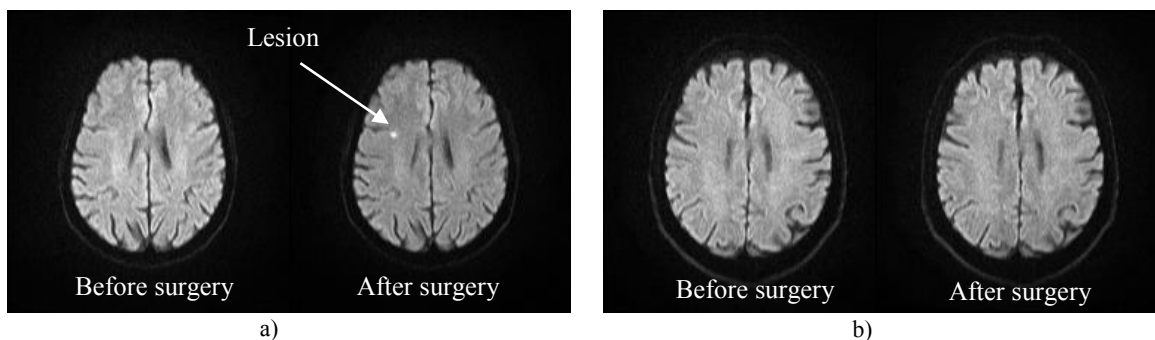

Figure 1. Cases of lesion detection from EMOCAR databases: (a) presence of new lesion and (b) no lesion

For instance, pulse sequences and hardware are often slightly different from a manufacturer to another which respectively impacts on the contrast and the noise of the image. Patient related artifacts include voluntary and involuntary motion during the course of scans but also mispositioning between scans. They are visualized on the images by translation or rotation distortions. It is obvious, that these kinds of artifacts may modify the shape and the localization of abnormalities.

Many tools have been developed over the last decades to detect lesions or tumors. The advantages and limits of medical image segmentation such as thresholding, pixels-based classification or region-based approaches segmentation are well-known[1]. Numerous of them use some priori knowledge regarding the shape or the location of the tumors. Furthermore, pixels' intensity in medical image may be quite similar between pathological and healthy tissues. Therefore, manual delineation is widely used but is time and resource consuming[2]. User interaction related is a key point in medical segmentation issue.

Our work aims to develop a standardized diagnostic support for subjective evaluation in order to reduce the risk of bias in clinical trials. Our key stake is to provide a detection tool mimicking the human visual system behavior during the visual inspection. The relevant HVS properties may be integrated in our lesion detection algorithm which must be robust to common distortions in medical images. Quality metrics are usually designed to predict the perceived quality of a natural scene image or video. The quality scores produced as output correlate to the Mean Opinion Score provided by the observers during subjective experiments. This key step of the quality evaluation scenario consists in recording the observers' quality rating of the images or videos.
Instead of predicting the quality, the metrics could be used to detect a difference between two images such as a lesion. Basically, the OQMs may be used to track a visibility threshold, i.e. the OQMs may tell us if there is a perceived difference between two given input images. A few previous studies have already been conducted on the possible use of OQMs in a tumor detection framework[3-5].

As a first step, the distortions commonly met during medical imaging, will be highlighted and their simulations process as well the lesion simulation method will be described. In the third section, we will present the results acquired during all the testing phase. To conclude, the ability of objective quality metrics to detect lesions on medical images will be discussed according the results obtained and the literature review.

2. MATERIALS AND METHODS

2.1. Distortion characterization

Our method was designed to understand deeply the behavior of objective quality metric to common distortions. Each single distortion identified as common in medical imaging, was tested independently on each volume. Both input volumes are different from lesion presence and distortion applied.

2.1.1. MRI variabilities and imaging distortions

Some authors already show that the quality of clinical diagnostic may result from operator-related factors such as manual lesions segmentation, or instrumentation-related factors with data-processing related factors[6, 7]. X. Xan, J. Jovicich et al. mention six noticeable MRI variability factors for cortical thickness measurement: modality manufacturer, pulse sequences, scanner field strength (1,5 T and 3T) and data processing factors. In the literature, it is widely accepted that the noise is related to the magnetic resonance system, the image processing and the physiological noise. For instance, final noise is related to the image voxel size, the receiver bandwidth and the number of averages in the image acquisition. The pulse sequence parameters (TR, TE) and the tissues properties (T1, T2, …) impact on the images contrast.
According the EMOCAR's protocol guidance, all data acquisitions are performed on identically configured diffusion MRI in pre- and post- surgery. Two pulse sequences are tested: FLAIR and Diffusion. Their parameters are respectively FOV 220 * 220, TE/TR 10 000/120 and FOV 230, TE/TR 84/3200.

2.1.2. Distortions simulation

According the variability factors detailed above, four kinds of distortions (grayscale, noise, rotation, scaling) has been simulated. The robustness of objectives quality metrics will be tested on the distorted images.

Table 1. Distortions generated to assess the quality metric robustness

DISTORTIONS	DESCRIPTIONS
Gaussian noise addition	The noise distribution for noisy MRI data is rician. However, when the signal-to-noise ratio is superior to two, the noise distribution can be approximated as Gaussian[8]. The noise in our volume was considered to have a Gaussian distribution. The amplitude of the simulated noise is not exceeded 10% of the maximum values of the slice.
Average grayscale modulation	For an extensive analysis, a large grayscale variation was considered, the pixels' intensity of each image were multiplied by the following coefficient: 0.98, 0.95, 0.93, 0.9, 0.88, 0.85, 0.83 and 0.8.
Rotation	Some authors indicate that a 1° rotation would be quite common in fMRI study. This value can be applied on MRI exams[10].
Scaling	A scaling distortion may be visualized on the output image related to the acquisition parameters selected. This kind of distortion can quite deeply reduce the quality metrics robustness.

Each single distortion was added one at the time on the volume. After applying all distortions on volumes from patient A and B, our dataset was composed of 203 images with scaled modifications, 174 images with rotated angles, 261 images with average grayscale modulation, and 174 images with Gaussian noise addition.

Furthermore, we need to test the limits of objective quality metrics by combining several kinds of distortions on a same image. Three distortions were therefore added successively on patient A: grayscale *0.8, noise with a standard deviation equal to 15% of the maximum value of each slice and 0.1° rotation.

2.1.3. Lesion simulation

The data from two patients, identified as A and B, are analyzed in this paper. For both, two input volumes are considered for each metric: the reference stack acquired before surgery and the distorted one which was generated by adding the lesions observed after surgery on the reference. Each volume is composed of 29 slices.
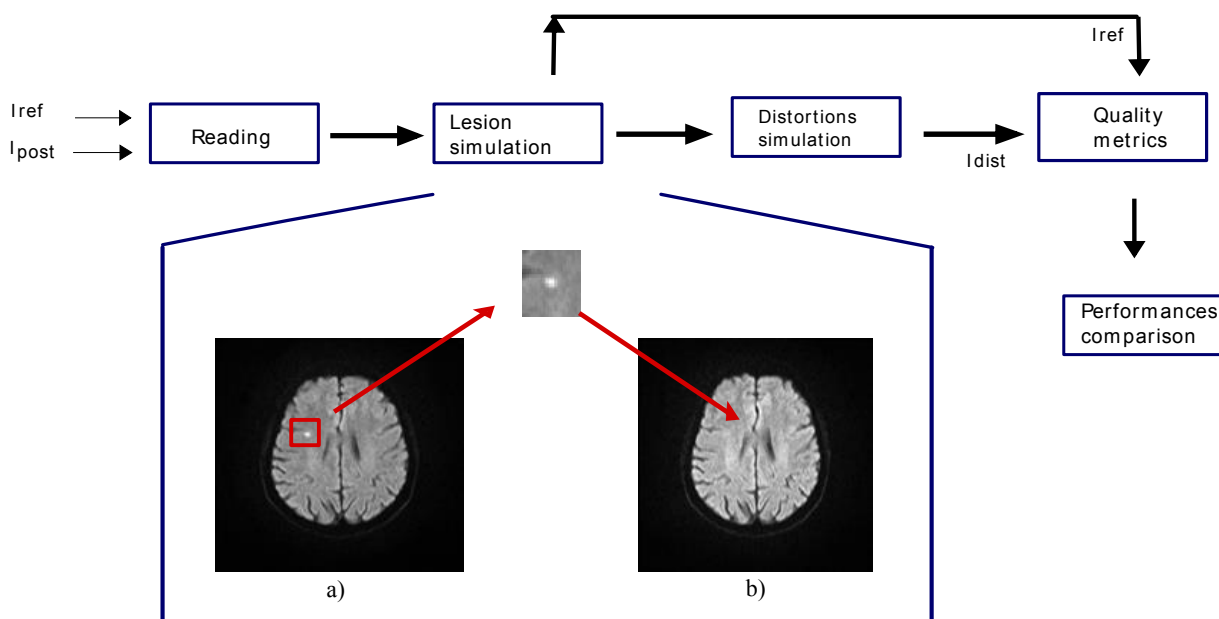

Figure 2. Lesion simulation process: (a) post-surgery image with lesion and (b) pre-surgery image with cropping lesion

Following this methodology, five lesions were simulated in distinct areas of the brain. Each lesion area has its own parameters in terms of contrast and localization with specific surrounding anatomical structures (Interhemispheric fissure,

lateral ventricle, sulcus, gyrus). Three lesions were added on patient A (slice 18, 19 and 23) and two on patient B (slice 8 and 14).

2.2. Objective Quality metrics

An objective quality metric (be it perceptual or statistical) can be of three kinds. Full reference metrics would take as inputs both the original image and its distorted version, a reduced reference metric would take as inputs some features of the original image, as well as the distorted image, finally, a no reference metric only uses the distorted image. Evidently, the metrics exhibiting the best performances are commonly the full reference metrics[9, 11].

In the context of EMOCAR, a full reference metric is the most adapted. A pre-surgery image is considered as the reference image and compared with the images under assessment (distorted ones). Eight quality metrics were tested on the four sets of distorted images in which the intensity of distortions fluctuate.

A few perceptual quality metrics have ever been designed specifically for medical images. Moreover, in the literature, few OQMs have been designed in order to withstand the distortions encountered in medical image acquisition. Indeed, very few metrics can overcome the geometrical distortions.

2.2.1. Spatial based quality metrics

The MSE (Mean Squared Error) and the related PSNR (Peak Signal Noise to Ratio) are the simplest and most widely used full-reference quality metrics. Both are based on the digital values of the pixels' intensity rather than on the physical luminance of distortion[12]. Many experiments showed that these tools do not correlate very well with perceived visual quality. PSNR does not take into account the human visual system properties.

Given the limitation of PSNR, Wang et al. propose the structural similarity (SSIM) index which is based on the ability of the HVS to deal with structural information. Despite these encouraging findings, this metrics has its limitations, for instance, it does not take into account the viewing conditions, i.e. viewing distance. Some optimization works have been performed to improve the metric performances[13, 14]. A multi-scale structural similarity method supplying more flexibility than previous single-scale one by incorporating the variations of viewing conditions were developed by Wang et al[14].

All these spatial based quality metrics such as PSNR and SSIM need a perfect co-registration between both input images. Their application onto medical images in our scenario may be compromised, as they basically compare the spatial difference between two images pixel by pixel or block per block for the PSNR and the SSIM respectively. For our purpose, the frequency information need to be considered in complementary of the spatial information.

2.2.2. Perception-based approaches to picture Quality Assessment

The VSNR is a wavelet-based visual signal-to-noise ratio. This metric quantifies the image fidelity of distorted image by considering the HVS properties. The VSNR considers both the contrast threshold detection model and the property of global precedence. It measures the degree to which the distortions disturb the image fidelity (structural and global features)[12].

The FSIM is based on the fact that the HVS understand an image mainly by analyzing its low-level features such as zero-crossing and edge. This metric uses two main parameters: the phase congruency (PC) and the Image Gradient Magnitude (GM). The gradient operator encodes the contrast information and is expressed by convolution masks. The main advantage of this metric is that the phase congruency (PC) is invariant to contrast fluctuation[15].

The VIF (Visual Information Fidelity) metric is based on SSIM and optimizes its performances by integrating the notion of mutual information. This is the ratio between two mutual information: the first is computed between the input and the output of the HVS channels when no distortion channel is present; the second one is the mutual information between the input of the distortion channel and its output from the HVS channel. The computational complexity of the VIF is much higher than SSIM or MS-SSIM[16].

The NQM[17] is based on Peli's contrast pyramid, this metric considers mainly the four following parameters without taking into account the orientation sensitivity of the HVS:

- Variation in contrast sensitivity with distance, image dimensions and spatial frequency;
- Variation in the local luminance mean;
- Contrast interaction between spatial frequencies;
- Contrast masking effects.

2.2.3. HVS Model-Based Perceptual Distortion Metrics

S. Daly developed the concept of HVS model based perceptual distortion metrics. This metric considers some of the HVS properties. Three distinct components of the Human Visual System are modeled: Amplitude non linearity, contrast sensitivity function and masking[18].

HVS-modeling involves the decreasing sensitivity of the CSF for higher spatial frequencies. The masking function explains why similar distortions are disturbing in certain regions of an image while they are hardly noticeable elsewhere (Figure 3).

In 1992, S. Daly developed a well-known image distortion metric in this field, the Visual Differences Predictor (VDP). Many authors try to improve this reference algorithm. The HDR-VDP version 2.2.1 (High-Dynamic Range Visual Differences Predictor) is an extension of the VDP. It integrates a custom CSF adapted to cover a large luminance range[19, 20]. The images used to perform this study have a bit-depth of 16 bits per pixels, the pixel's intensity range is from 0 to 820. As output, the HDR-VDP version 2.2.1 provides a probability map, where the perceptual differences have a high probability to be visualized.
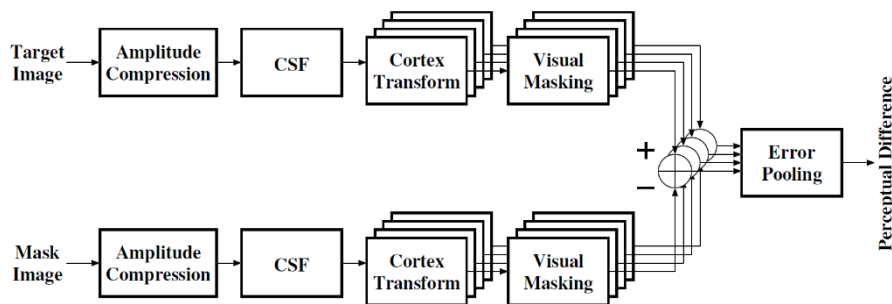

Figure 3. Data flow diagram of the Visual Differences Predictor (VDP)[19]

3. RESULTS

The performances of eight quality metrics were tested on medical images. The lesions were visible on slice 8 and 14 of the volume B and slice 18, 19 and 23 of the volume A. Four kinds of distortions (noise, grayscale and scaling distortion, and rotation) were added on the volume with lesions. For each distortion type, several versions were generated.

The metric scores for each quality metrics are plotted from Figure 4 to Figure 9. A lesion is identified by the metrics as a difference between the reference image and the distorted one, the metric scores should decrease accordingly. For all graphs, the quality scores provided by HDR-VDP, were divided by 2 for readability.

In this work, we first evaluated the metrics performance independently according to the various induced distortions (section 3.1), and then, in order to evaluate the global performances of the metrics and their ability to detect the lesion in a real-case scenario, we have grouped several distortions altogether (section 3.2).

3.1. Independent distortions evaluation

The results from the noise distortion test are plotted in Figure 4 and Figure 5. The metric based on structural similarity (SSIM) does not detect any differences on slices 18, 19 nor 23 for the patient A. The same applies for the patient B on slices 8 and 14. This metric compares the luminance, the contrast and the structure of the inputs images (reference and distorted volumes)[21]. In medical images, the contrast between a lesion and the healthy tissues are often low. The abnormalities are not also totally structurally defined. Some other areas such as the skull or the ventricles have stronger delimiting outlines. The main drawback of the SSIM is its inability to take into account the viewing conditions of the

experiment. The MS-SSIM show a better result with a small metric score decrease in slices where lesions are visible. From a standard deviation equal to 8 % of the maximum intensity of the slice, the MS-SSIM does not detect any lesion.

Moreover, in the same figures, PSNR values from slices to slices oscillate a lot. The Gaussian noise added on the images directly impacts their pixel intensities. The difference generated is thus applied to the whole images.

By analyzing these graphs, the perceptual quality metrics, VIF, FSIM, VSNR, HDR-VDP seem to give a better approximation of the localization of the lesion in the volumes. For the volume B, the metrics, HDRVDP and VIF, manage to detect the lesion in slice 14. This lesion was deliberately positioned in an area close to the lateral ventricle with a low contrast to simulate the worst conditions of detection. Such a behaviour is noticeable on the detection probabilities provided by the HDR-VDP metric, i.e. lower probability of detection on slice 14 than for slice 8 (Figure 6).
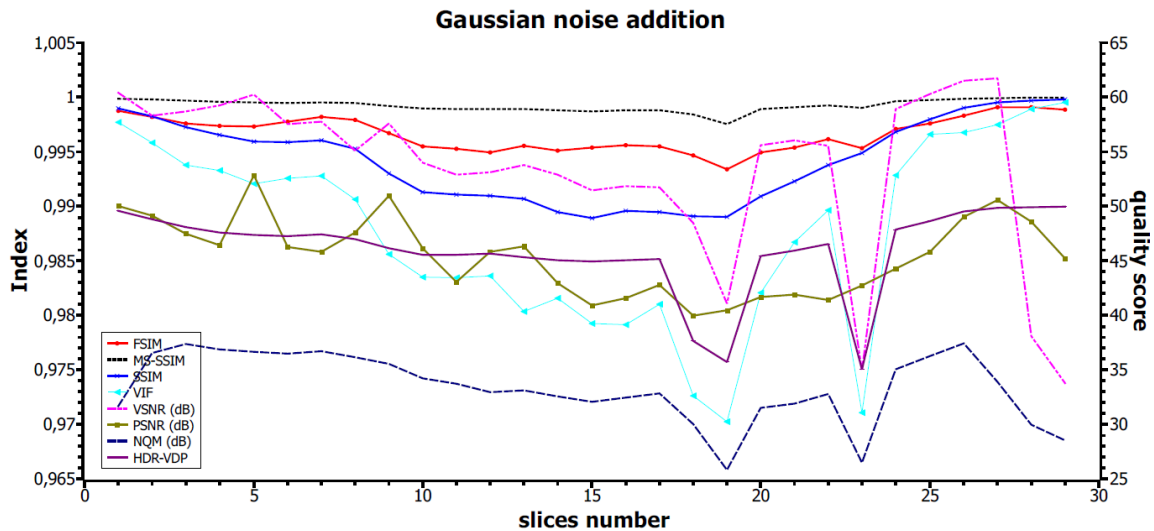

Figure 4. Gaussian noise addition generated on the volume A with a standard deviation equal to 5% of the maximum value of each slice
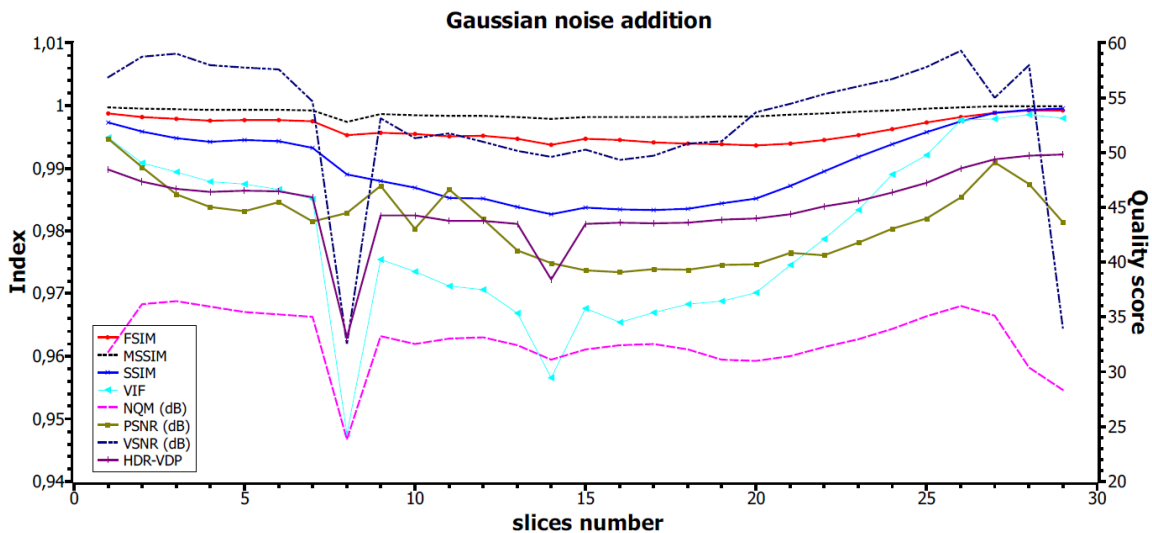

Figure 5. Gaussian noise addition generated on the volume B with a standard deviation equal to 5% of the maximum value of each slice
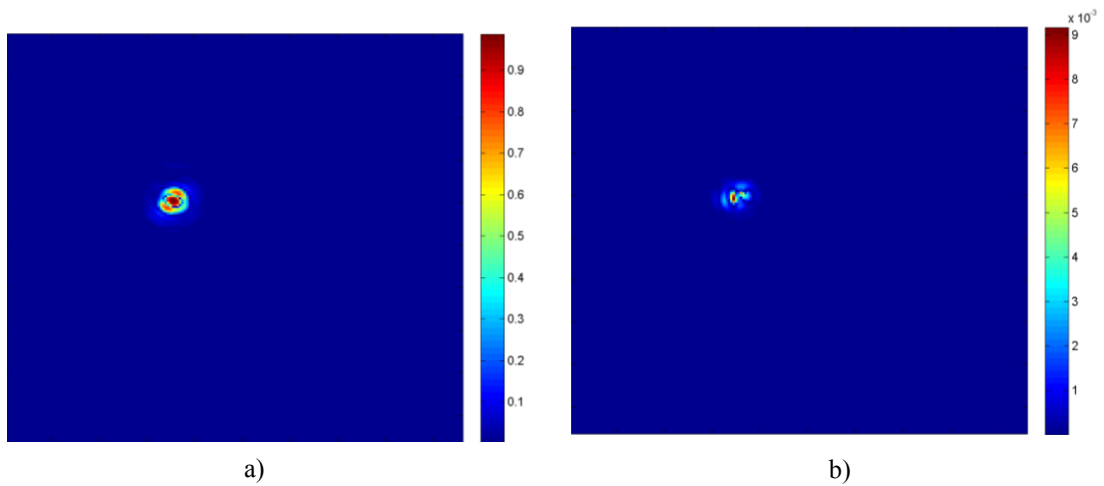
Figure 6. Probability maps of the HDRVDP produced on the volume B with a standard deviation equal to 5% of the maximum value of each slice: (a) slice 8 and (b) slice 14

The results from average grayscale modulation confirm the observations detailed above, the perceptual quality metrics offers a more accurate lesion detection. The outputs from the VSNR predict two eventual differences on the volume A in slices 23 and 28. The lesion in slice 28 can be considered as a false positive.

In Figure 7, the HDR-VDP seems to deal with a rotation of 0.5°. The lesion in slice 18, 19 and 23 for patient A and in slice 8 for patient B were detected. Unfortunately, by adding this kind of distortion, false detections may occur (see Figure 8).
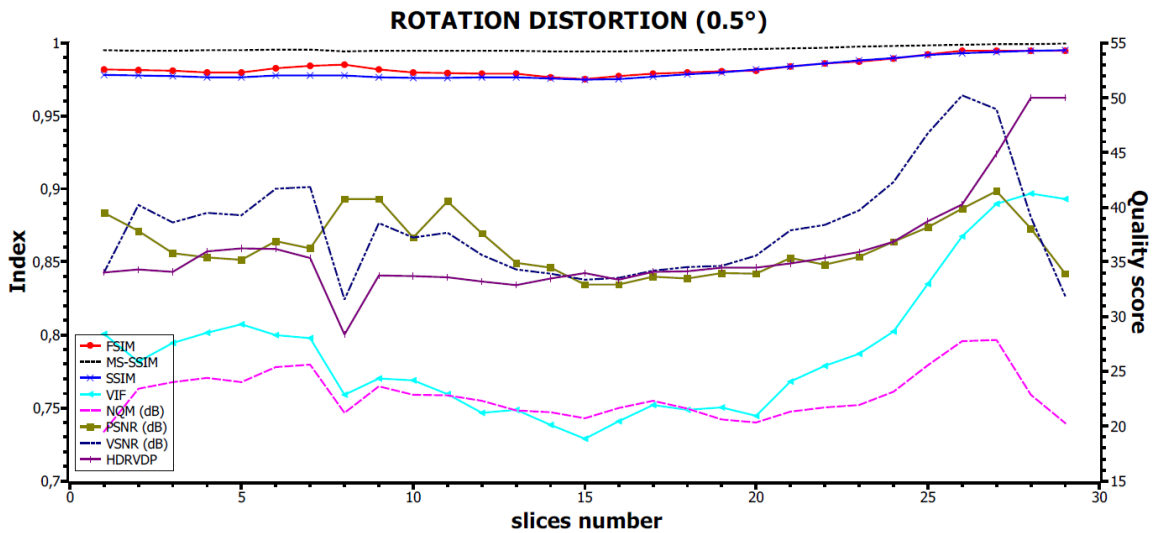

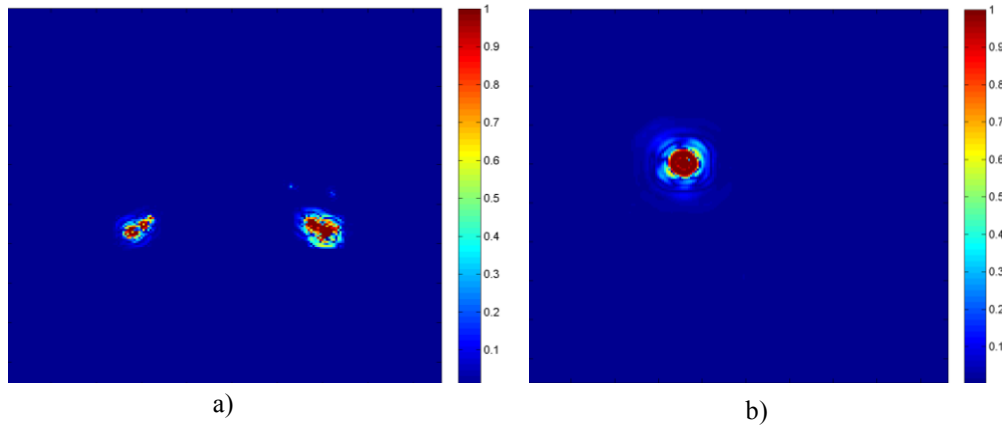Figure 7. Rotation distortion applied on the volume B with an angle of 0.5°

Figure 8. Probability maps produced by the HDR-VDP on the volume A with a rotation of 0.8°: (a) in slice 9 and (b) in slice 19

The metric scores show no detection of any lesion on both volumes for a low scaling distortion. All metrics seem to be very sensitive to spatial shifting, which have little effects on visual fidelity.

3.2. Multiple distortions

In section 2.1.2, we explained that it was relevant to superimpose several distortions in the same volume to approximate clinical cases. The metrics' scores from the volume A after applying a rotation, a grayscale modification and a Gaussian noise are plotted in Figure 9. No metric, excepted the HDR-VDP, provides a good approximation of the lesion detection. The quality scoring of this metric decreases for five slices: 5, 9, 18, 19 and 23. We notice the same false positive issue as described in previous section.
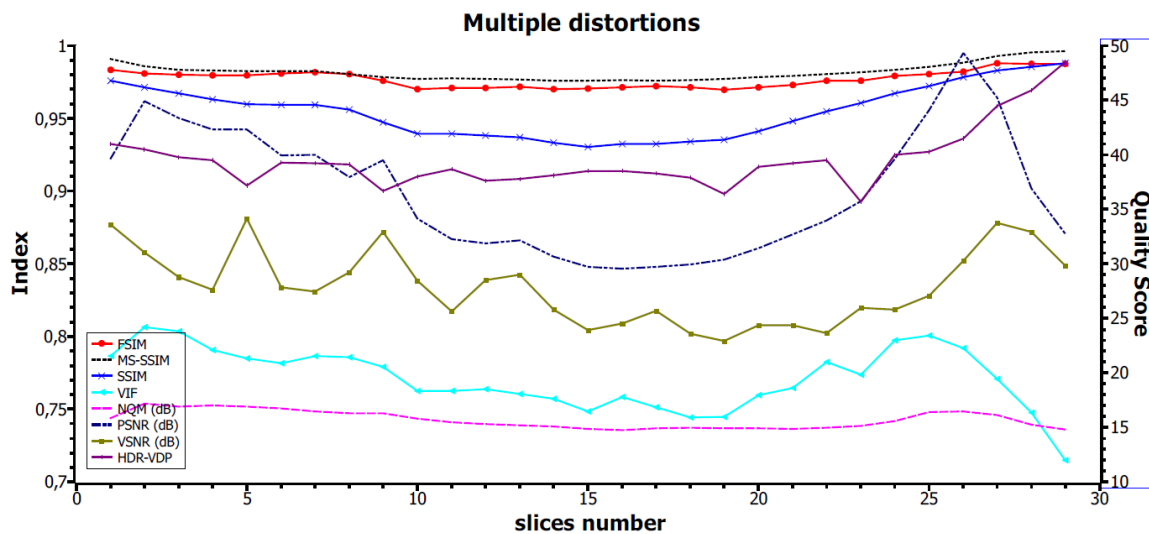

Figure 9. Multiple distortions applied on the volume A with a rotation of 0.1°, a grayscale modification of 0,8 and a Gaussian noise of 15% of the maximum value of each slice

4. DISCUSSION

In previous section, we showed that the perceptual quality metrics (VIF and HDR-VDP) provide a better estimation of the presence of a new lesion or shape distortion in the vascular tree. Although the MS-SSIM integrates a multi scale dimension

to the SSIM, its deficiency remains the same. Neither MS-SSIM, nor SSIM consider the grouping of errors, which may lead to a perceived quality loss, and hence to a visual detection.

The VIF metric shows some encouraging results. The interesting point is that this metric decomposes the image into sub-bands with different weights at the pooling stage. However, within the sub-band, each position has the same importance. To model the visual inspection, appropriate spatially varying weights should be integrated in the metrics.

The first stage of the VSNR is based on the analysis of the contrast threshold of the images which allows to determine if the distortion disrupts or not the global precedence. In our case, the contrast of the volume was modified by adding distortions. The initial step of the VSNR approach is compromised.

The pooling strategy performed by the FSIM does not consider the spatial aspect of the differences, the similarity value computed at the output is the product of the similarity of each parameter (the phase congruency and the gradient magnitude).

In section 2.2.3, we explained that the HDR-VDP integrates several HVS properties. This metric simulates the loss of the sensitivity of the human eye at high and very low frequencies using a contrast sensitivity function. The Gaussian noise added in the experiment is a high frequency noise, the CSF should filter an important part of this distortion. The initial step of the HDR-VDP metric being a grey level to perceived luminance conversion, a first normalization is thus brought onto the images. Such a process may lessen the grayscale difference between the input images. Moreover, the consideration of the masking effect and the decomposition into spatial and orientation channels allows the metric to consider the frequency selectivity of the visual system[18]. In section 3.1, we have demonstrated that the HDR-VDP was able to detect the lesion when Gaussian noise was added or the image grayscale amplitude was modulated. In Figure 4, the metric score decreases in slice 23 further than for slices 18 and 19. This observation does not correlate with the ground truth, the lesion in slice 19 being more visible. The lesion in slice 23 is located at the edge of the brain, where the contrast is very high (transition from the skull to the background image). In these conditions, a lesion less visible may be more detectable if it is positioned in a high contrast area.

In the next steps, some tools should be integrated or developed to determine a detection threshold method to meet the validation expectations for our detection method. Some clustering methods are already available in the literature, such as the dendrogram[22]. Figure 10 shows the dendrogram plotted with our objective data, three outliers are visible in our data obtained after addition of Gaussian noise. The slices index is shown on the x-axis, whereas the y-axis represents the distance between the clusters. The performance of this technique was not reproduced for all test situations. Some additional work is therefore required to adapt this clustering method to our current purpose.

The final expectation is to be able to use an adaption of the HDR-VDP on ground truth medical images. All our results were obtained from volumes with simulated lesions and known distortion level. In clinical conditions, the overall distortions the images may go through can significantly differ from our simulations (as presented in section 2.1.2). We have thus hereby tested the HDR-VDP metric on the actual images from the EMOCAR study. The before/after surgery images were acquired a few days apart, and their differences appear to be stronger than our modeled datasets. Figure 11 shows the metric scores on the actual EMOCAR images, as can be seen on this plot, none of the tested metrics is able to detect the lesion. The probability maps related for the slice 9 and 19, are given in Fig. 12. The differences in very bright area (spots in slice 9) are overestimated, this observation correlates with the previous facilitation process previously described. The high probabilities of differences along the skull may be related to the lack of registration between both volumes. In the preliminary stage of our algorithm, a registration would be needed, even if the HDR-VDP have shown some encouraging abilities to deal with the rotation distortions. Further studies will be devoted to a better consideration of the distortions, some modifications should be brought onto the HDR-VDP metric, in order to make it invariant to geometric distortion.
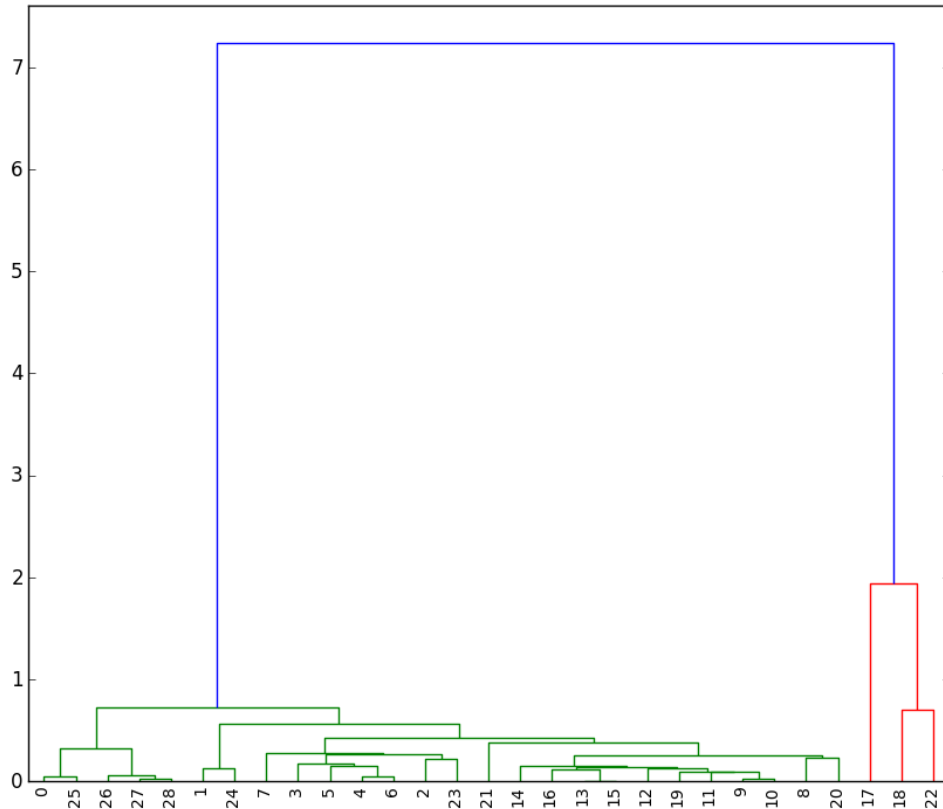
Figure 10 Dendrogram obtained with the VIF and the HDR-VDP data as inputs. These data were recorded on volume A with a Gaussian noise addition generated (standard deviation equal to 5% of the maximum value of each slice)
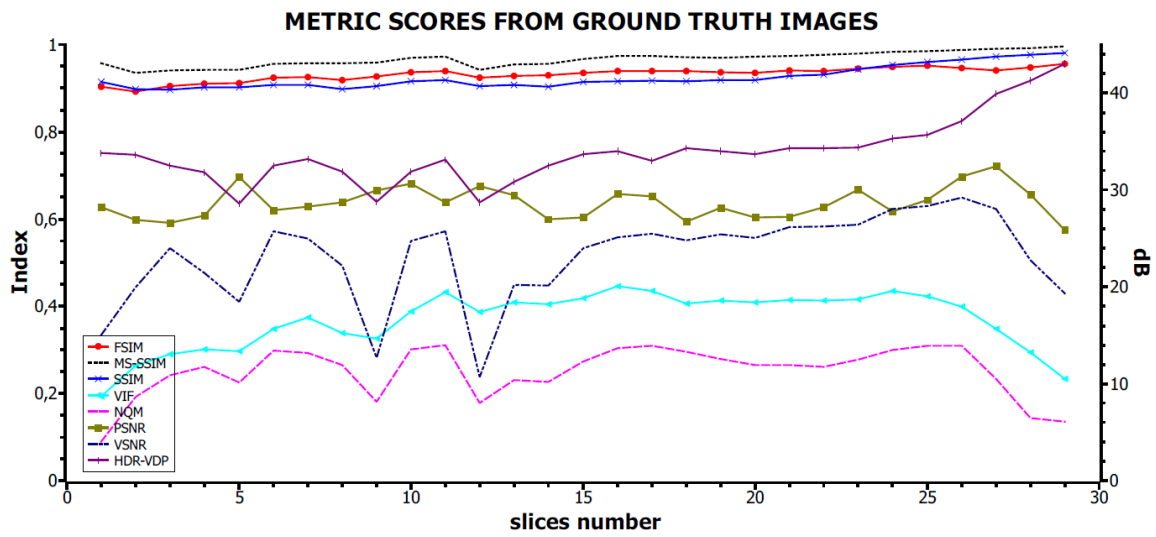

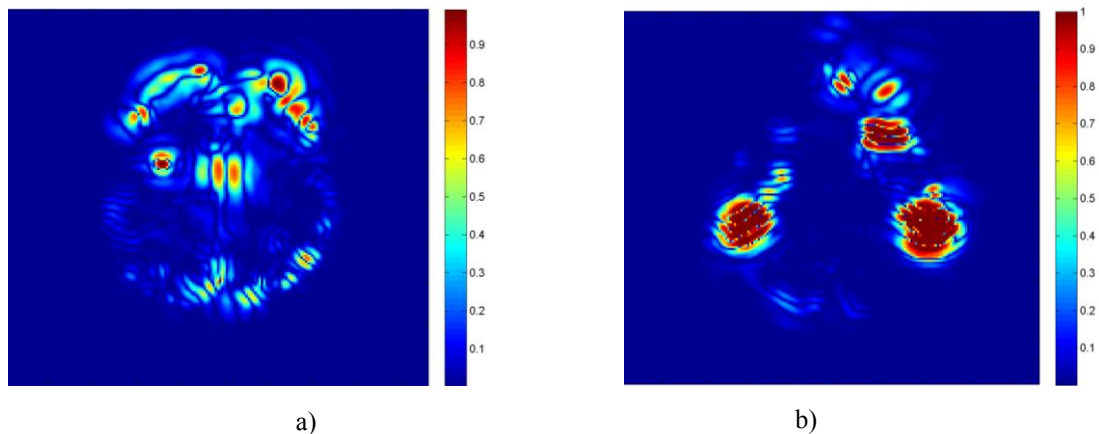Figure 11 Application of the HDR-VDP on ground truth images (patient A)

Figure 12 Probability maps produced by the HDRVDP on the images before and after surgery: (a) in slice 19 and (b) in slice 9

5. CONCLUSION

The perceptual metrics such as HDR-VDP and VIF have shown in this paper some abilities to detect vascular lesions on MR images. By converting the images into perceived luminance, and filtering with the CSF, the HDR-VDP have demonstrated its robustness to contrast and grayscale distortions. However, some limitations have been identified, the masking effect typical of this metric, involves an overestimation of the differences in the high contrast area.

Some additional works should be done to improve the robustness of this metric to geometrical distortions. Preliminary steps such as a registration process could improve the performance of the HDR-VDP for detecting new lesions or shape distortion into the vascular tree.

Next steps will also focus on the development of an optimized version of the dendrogram. Indeed, all the analyzed metrics, except the PSNR and the SSIM, show a more or less significant decrease of quality scores for the slices where lesions are visible. An adapted threshold detection method could be used to differentiate real lesions area from normal tissues.

Finally, the cost risk-benefits for the clinical trial will need discussed both for the patient and for the expert.

REFERENCES

[1] Petitjean, C., and Dacher, J. N., "A review of segmentation methods in short axis cardiac MR images," Medical image analysis, 15(2), 169-184 (2011).

[2] Shen, S., Szameitat, A. J., & Sterr, A., "Detection of infarct lesions from single MRI modality using inconsistency between voxel intensity and spatial location—a 3-D automatic approach," IEEE Transactions on Information Technology in Biomedicine, 12(4), 532-540 (2008).

[3] L. Zhang, C. Cavaro-Ménard, P. Le Callet, D. GE.,"A multi-slice model observer for medical image quality assessment," IEEE International Conference on ICASSP, 1667-1671 (2015).

[4] Zhang, L. et al. "A perceptually relevant channelized joint observer (pcjo) for the detection-localization of parametric signals," IEEE Transactions on Medical Imaging, 31(10), 1875-1888 (2012).

[5] Jackson, W. B., Said, M. R., Jared, D. A., Larimer, J. O., Gille, J. and Lubin, J., "Evaluation of human vision models for predicting human observer performance," Medical Imaging 1997, 64-73 (1997). International Society for Optics and Photonics.

[6] Han, X., Jovicich, J., Salat, D., et al. "Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer," Neuroimage, 32(1), 180-194 (2006).

[7] Jovicich, J., Czanner, S., Greve, D., et al. "Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data," Neuroimage, 30(2), 436-443 (2006).

[8] Gudbjartsson, H., & Patz, S., "The Rician distribution of noisy MRI data," Magnetic resonance in medicine, 34(6), 910-914 (1995).

[9] Zhang, J., & Le, T. M. "A new no-reference quality metric for JPEG2000 images," IEEE Transactions on Consumer Electronics, 56(2), 743-750 (2010).

[10] Jezzard, P., & Clare, S. "Sources of distortion in functional MRI data," Human brain mapping, 8(2-3), 80-85 (1999).

[11] Angelo, A. D., Zhaoping, L., & Barni, M., "A full-reference quality metric for geometrically distorted images," IEEE Transactions on Image Processing, 19(4), 867-881(2010).

[12] Chandler, D. M., & Hemami, S. S. "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," IEEE Transactions on Image Processing, 16(9), 2284-2298 (2007).

[13] Chai, L., Sheng, Y., & Zhang, J., "SSIM performance limitation of linear equalizers," IEEE International Conference on ICASSP, 1220-1224 (2014).

[14] Wang, Z., Simoncelli, E. P., & Bovik, A. C., "Multiscale structural similarity for image quality assessment," Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, Vol. 2, 1398-1402 (2003).

[15] Zhang, L., Zhang, L., Mou, X., & Zhang, D., "FSIM: a feature similarity index for image quality assessment," IEEE Transactions on Image Processing, 20(8), 2378-2386 (2011).

[16] Sheikh, H. R., & Bovik, A. C. "Image information and visual quality," IEEE Transactions on Image Processing, 15(2), 430-444 (2006).

[17] Damera-Venkata, N., Kite, T. D., Geisler, W. S., Evans, B. L., & Bovik, A. C., "Image quality assessment based on a degradation model," IEEE Transactions on Image Processing, 9(4), 636-650 (2000).

[18] Daly, S. J. "Visible differences predictor: an algorithm for the assessment of image fidelity," In SPIE/IS&T Symposium on Electronic Imaging: Science and Technology (pp. 2-15). International Society for Optics and Photonics (1992).

[19] Mantiuk, R., Myszkowski, K., & Seidel, H. P., "Visible difference predicator for high dynamic range images," IEEE International Conference on Systems, Man and Cybernetics, 3, 2763-2769 (2004).

[20] Mantiuk, R., Kim, K. J., Rempel, A. G., & Heidrich, W, "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," In ACM Transactions on Graphics, 30(4), 40 (2011).

[21] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, 13(4), 600-612 (2004).

[22] Phipps, J. B., "Dendrogram topology," Systematic Biology, 20(3), 306-308 (1971).